

## En la Universidad de Stanford

# Joven costarricense combina investigación en genómica y biofísica

El 2014 fue un buen año en materia de publicaciones para el Dr. Carlos Luis Araya Pereira, investigador del Departamento de Genética de la Universidad de Stanford, y ex-alumno de ingeniería en biotecnología del Instituto Tecnológico de Costa Rica (TEC). El académico publicó tres artículos en *Nature* y *Nature Biotechnology*.

Como miembro de dos laboratorios distintos (Michael Snyder y William Greenleaf) en la universidad mencionada, este doctor en genómica desarrolla investigación científica que abarca temas en biofísica, evolución, genómica y, actualmente, cáncer.

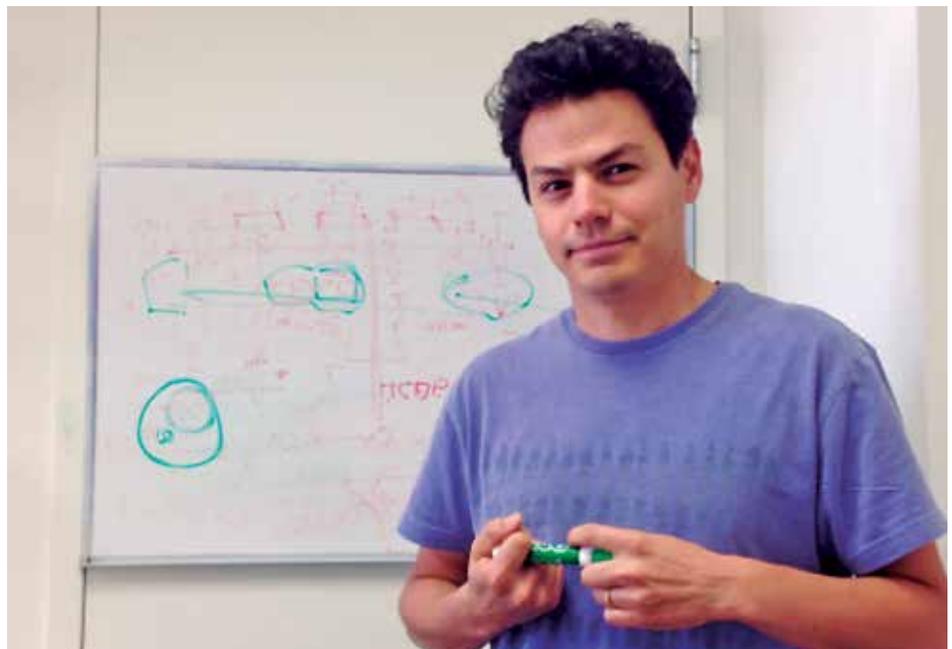
Los artículos se pueden revisar en los siguientes enlaces:

<http://www.nature.com/nbt/journal/v32/n6/full/nbt.2880.html>

<http://www.nature.com/nature/journal/v512/n7515/full/nature13497.html>

<http://www.nature.com/nature/journal/v512/n7515/full/nature13668.html>

Los artículos más recientes (Araya et al.[1] y Boyle et al.[2]) –desarrollados como parte de ENCODE/modENCODE (<http://www.genome.gov/encode/>)– analizan la actividad de las proteínas que regulan la síntesis de genes. En Boyle et al.[2], el equipo evaluó la localización de más de 300 proteínas regulatorias en el genoma de células humanas y dos de los principales modelos multicelulares, el nemátodo *C. elegans* y el díptero *D. melanogaster*. La actividad de estos reguladores en diferentes contextos a lo largo de 1019 experimentos (de ChIP-seq; inmuno-precipitación de cromatina y secuenciamiento) les permitió evaluar una multitud de aspectos de circuitos de regulación genómica y la conservación de dichas propiedades a largas distancias evolutivas (el



Dr. Carlos Araya Pereira.

urbilaterano, el último ancestro en común entre las tres especies, se estima que vivió hace cerca de 570 millones de años.)

Según expresó el doctor Araya, el artículo Araya et al.[1] representa un nuevo enfoque: la integración de información sobre la localización genómica de proteínas regulatorias con experimentos que identifican las células específicas y fases de desarrollo embrionario en las que dichas proteínas son activas. La estrategia permitió, por vez primera, analizar en paralelo los linajes específicos en que cada regulador está activo y los genes que regula.

Estos trabajos fueron objeto de referencia reciente en *Nature News* (<http://www.nature.com/nature/journal/v512/n7515/full/nature13668.html>) y el *New York Times* (<http://www.nytimes.com/2014/09/02/science/human-fly-worm-dna.html>).

Sin embargo, según el científico, el estudio de mayor impacto fue publicado en *Nature Biotechnology*. Para comprender su significado es necesario un poco de contexto tecnológico.

Durante la última década, la eficiencia de tecnologías de secuenciamiento de ADN se ha duplicado anualmente. Por ejemplo, en cuanto en el 2001 costaba cerca de tres mil millones de dólares obtener la secuencia de un genoma humano, hoy, lo mismo tiene un costo aproximado de mil dólares.

Carlos Luis Araya explica que a los ingenieros en computación les gusta hablar de la erróneamente llamada Ley de Moore (que en realidad

es una observación, no una ley) como ejemplo de la tasa de avance tecnológico. Dicha observación se refiere a la tasa de avances en procesadores, cuya densidad de circuitos (y aproximada capacidad computacional) se ha duplicado más o menos cada 18 meses por más de 40 años.

Sin embargo, los avances en el secuenciamiento de ADN son aproximadamente un 50% más rápidos. Dichas plataformas tecnológicas integran avances en biología molecular, enzimología, física óptica, nano-manufactura e ingeniería de computación, para secuenciar  $>10^9$  moléculas de ADN en paralelo. Dichos avances son el corazón de la llamada revolución genómica. No obstante, recalca el doctor Araya, a pesar de todos sus logros, dicha revolución por sí sola no nos permite comprender el funcionamiento de las moléculas codificadas en genomas.

Explica que los ensayos de actividad bioquímica y biofísica (como afinidad, actividad cinética y actividad catalítica) no han tenido los mismos incrementos masivos en escala. Por ejemplo, las tecnologías comerciales existentes permiten realizar cerca de mil ensayos bioquímicos de afinidad entre proteínas u otras biomoléculas. Es decir, ensayos de bioquímica y biofísica están limitados a escalas que son  $10^6$  veces menor que las escalas alcanzadas por tecnologías de secuenciamiento de ADN.

Con esto en mente, el equipo del laboratorio de William Greenleaf hackeó [3] un instru-



mento de secuenciación moderno para, possecuenciación, sintetizar más de  $10^9$  moléculas de ARN *in situ* y analizar cuantitativamente su afinidad por proteínas.

El doctor Araya añade que “la nueva plataforma tecnológica nos permitió realizar el mayor experimento de afinidad bioquímica a la fecha, con más de  $10^7$  mediciones energéticas de afinidad y más de  $10^3$  mediciones cinéticas entre moléculas de ARN y proteínas. En un solo experimento de afinidad entre componentes del virus MS2, pudimos reconstruir más de cuatro décadas de investigación bioquímica en el modelo clásico de interacción entre ARN y proteínas. El volumen y la calidad de los datos nos permitieron construir un modelo predictivo de la afinidad y evaluar las consecuencias evolutivas de distintas propiedades biofísicas del ARN”. Este último punto, recalca el doctor Araya, implica que es posible predecir aspectos de trayectorias evolutivas de biomoléculas cuya evolución está rigurosamente restringida por propiedades biofísicas.

### Formación

Carlos Luis Araya Pereira estudió en la escuela pública de Tres Ríos y realizó sus estudios secundarios en el Colegio de San Luis Gonzaga, en Cartago. Posteriormente entró al TEC a estudiar biotecnología, donde estuvo dos años hasta que en el 2001 viajó a los Estados Uni-

dos para continuar sus estudios.

Este joven investigador terminó su carrera de biotecnología en la *Washington State University*, donde completó con honores un bachillerato en genética y biología celular y un bachillerato en biotecnología, en el 2003. Posteriormente fue invitado al programa de doctorado en el Departamento de Ciencias Genómicas en la Universidad de Washington, en Seattle y, desde el 2011, realiza estudios posdoctorales en el Departamento de Genética de la Universidad de Stanford.

En la Universidad de Washington, junto con Stanley Fields y Douglas Fowler, inventó tecnologías [4,5] para el análisis de millones de variantes de proteínas en paralelo. Según el doctor Araya, una de sus mayores satisfacciones ha sido ver cómo estas tecnologías han sido adoptadas por investigadores y han transformado el estudio de proteínas. No obstante, recalca que esta primera generación de tecnologías no provee información energética o catalítica de la actividad de biomoléculas, sino una aproximación de ellas. Las nuevas tecnologías –que han sido llamadas un *tour de force* tecnológico por su integración de métodos de genómica, biología molecular y biofísica– prometen una verdadera revolución bioquímica.

1. Araya CL, Kawli T, Kundaje A, Jiang L, Wu B, et al. (2014). Regulatory analysis of

the *C. elegans* genome with spatiotemporal resolution. *Nature* 512: 400–405.

2. Boyle AP\*, Araya CL\*, Brdlik C, Cayting P, Cheng C, et al. (2014). Comparative analysis of regulatory information and circuits across distant species. *Nature* 512: 453–456.
3. Buenrostro JD\*, Araya CL\*, Chircus LM, Layton CJ, Chang HY, et al. (2014). Quantitative analysis of RNA-protein interactions on a massively parallel array reveals biophysical and evolutionary landscapes. *Nature Biotechnol.* Available: <http://dx.doi.org/10.1038/nbt.2880>.
4. Fowler DM, Araya CL, Fleishman SJ, Kellogg EH, Stephany JJ, et al. (2010). High-resolution mapping of protein sequence-function relationships. *Nature Methods* 7: 741–746.
5. Araya CL\*, Fowler DM\*, Chen W, Muniez I, Kelly JW, et al. (2012). A fundamental protein property, thermodynamic stability, revealed solely from large-scale measurements of protein function. *Proceedings of the National Academy of Sciences* 109: 16858–16863. ■